A complex network diagram with numerous nodes of varying sizes (dark blue, light blue, and grey) connected by thin grey lines. Some nodes are highlighted with larger concentric circles. The background is white with faint, larger-scale network patterns.

ACQUAINTANCE NETWORK OF THE MARVEL CINEMATIC UNIVERSE

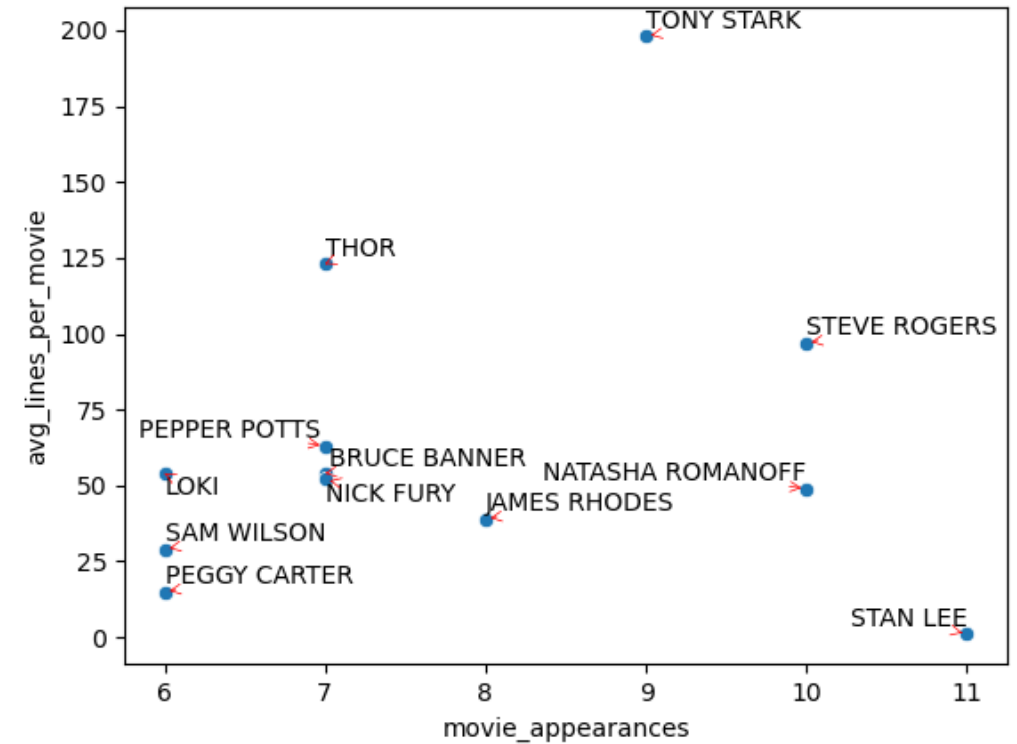
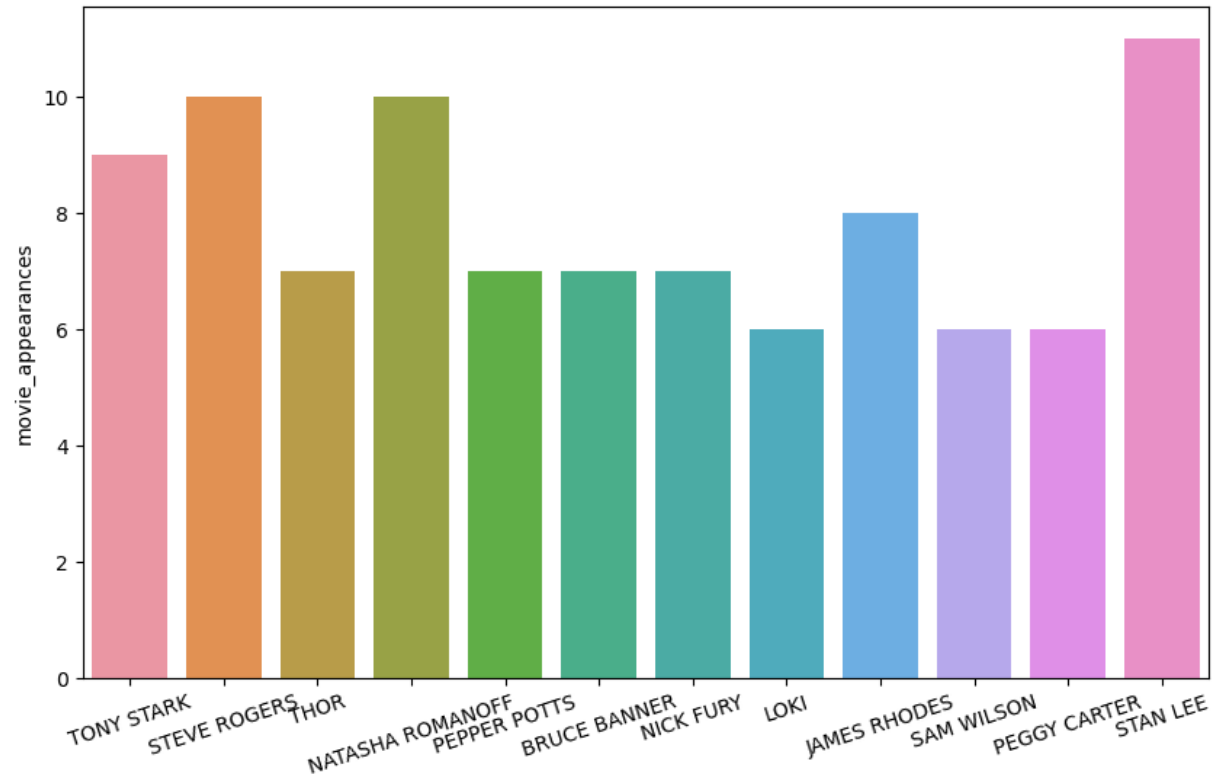
Dajka Bence

GOAL

- Obtain the acquaintance network of the characters in the MCU and calculate some basic characteristics.
- Tasks:
 - Create the network from the data
 - Visualize it
 - Calculate: degree distribution, clustering coefficient, degree correlation

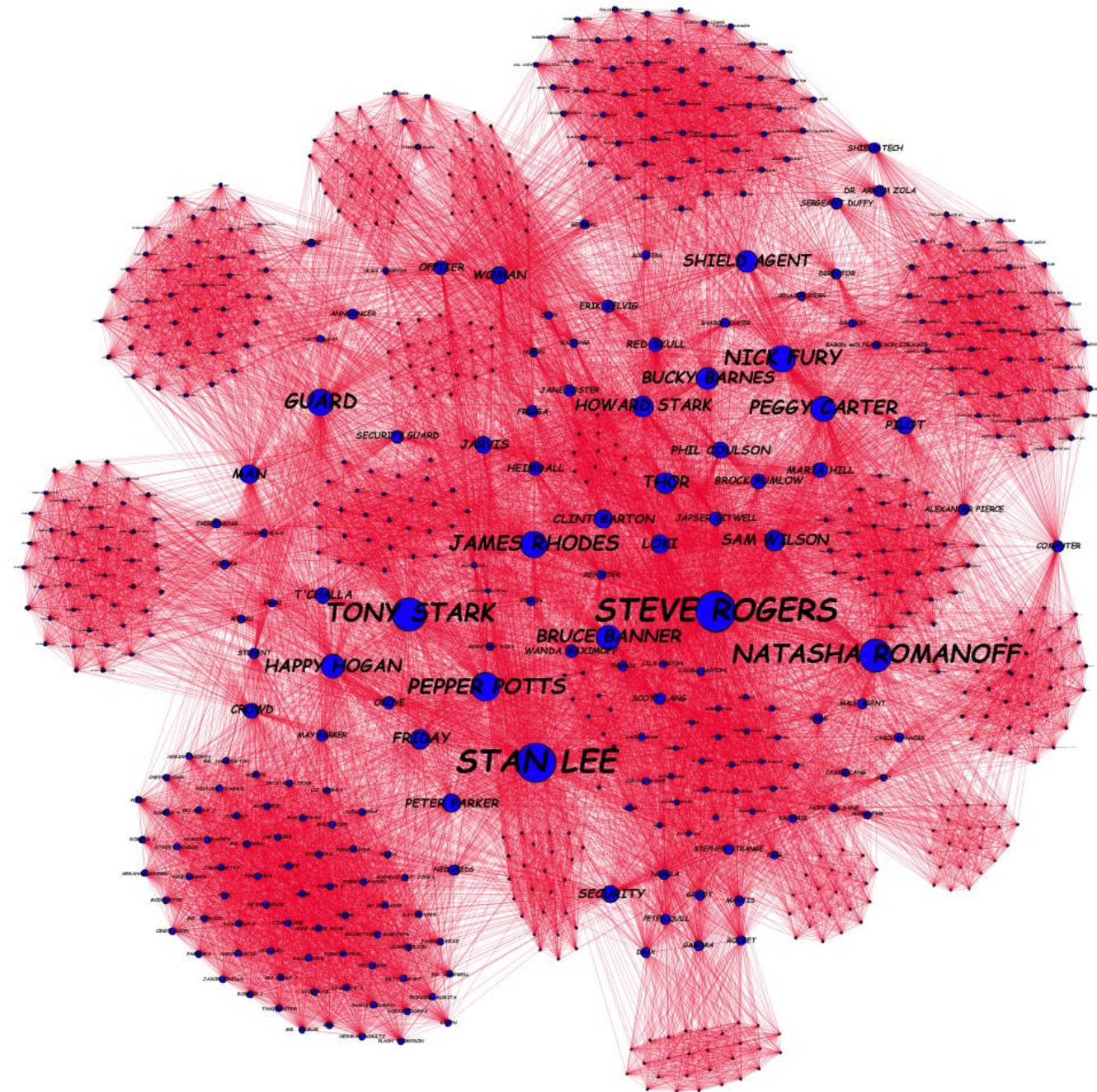
DATA

- Three tables: movies, characters, mcu
- Create cast table
- Add UUID to the characters



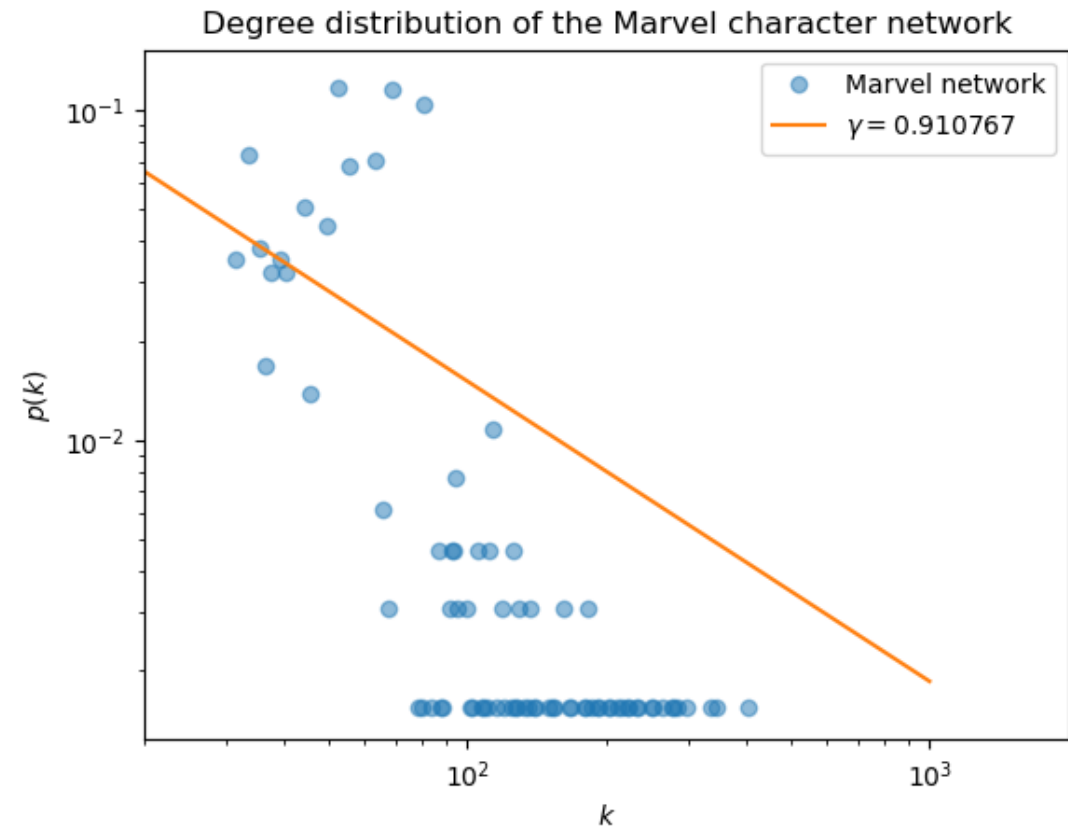
THE NETWORK

- Stan Lee: creator of the characters
- Steve Rodgers: **Captain America**
- Tony Stark: **Iron Man**
- Natasha Romanoff: **Black Widow**



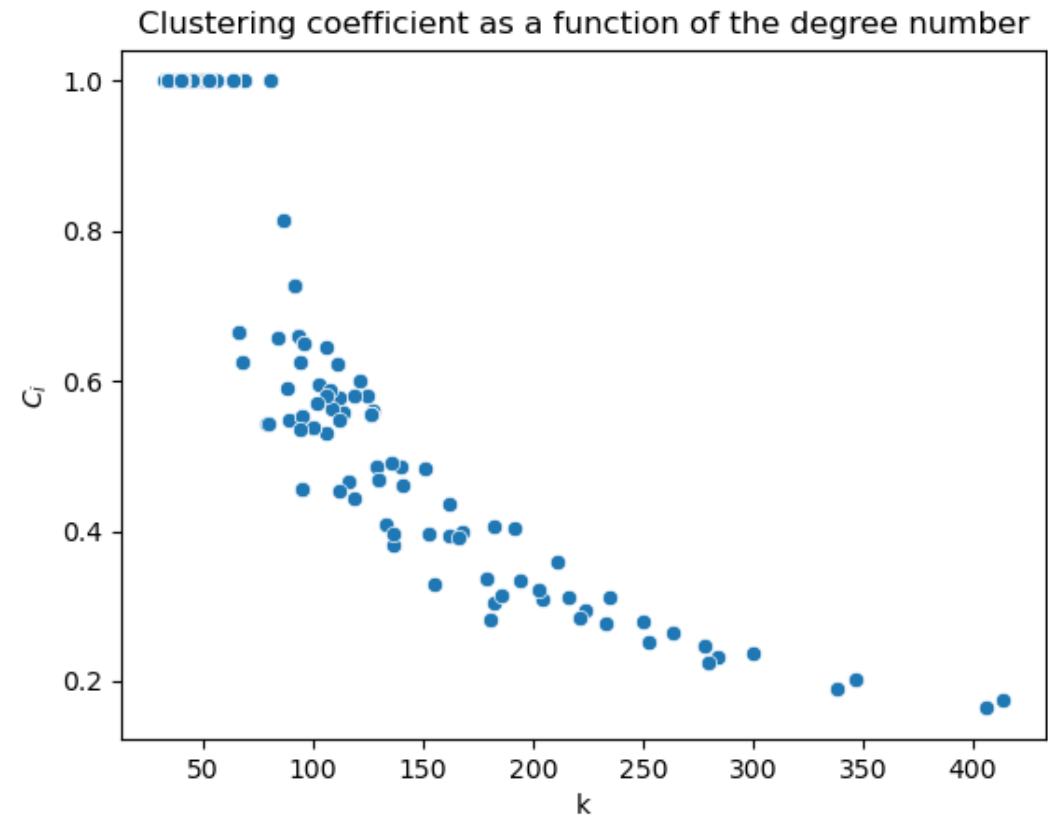
DEGREE DISTRIBUTION

- Node degree represents the number of links one node has to other nodes.
- $\langle k \rangle = 67.93$
- scale-free property: $p(k) \sim k^{-\gamma}$



CLUSTERING COEFFICIENT

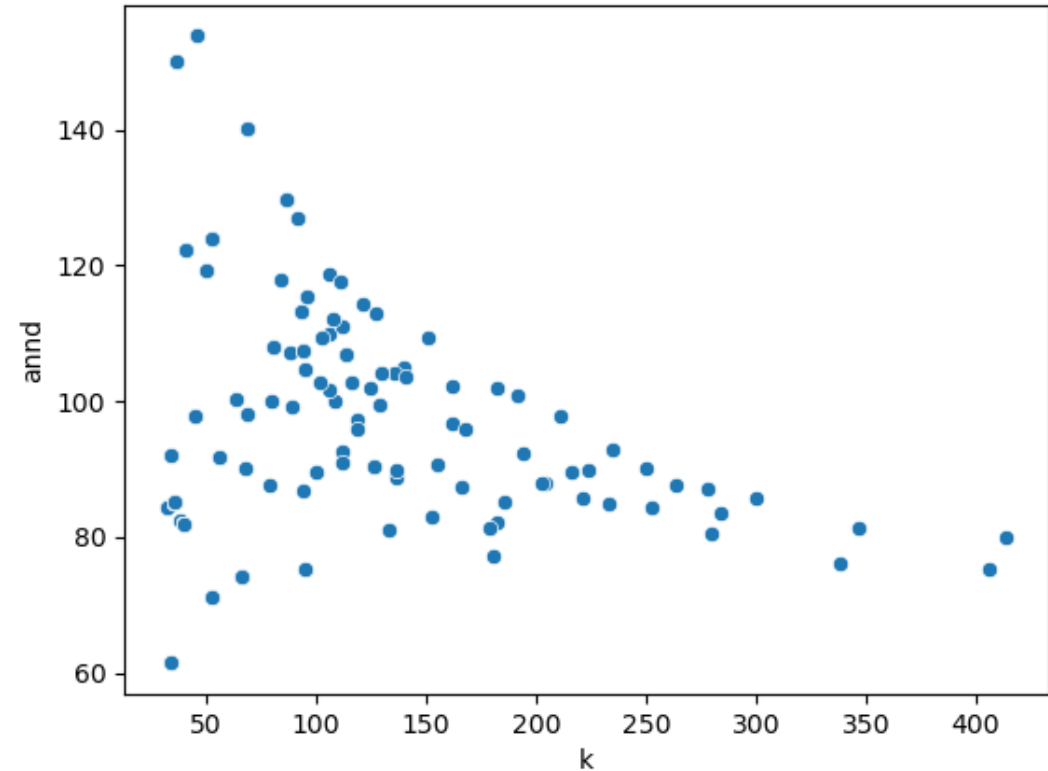
- Measures how many of the characters neighbors (co-actors) are neighbors with each other.
- C_i is the probability of the neighbors linking to each other.
- $\langle C \rangle = 0.92$
- The global clustering coefficient CC measures the total number of closed triangles in the network.
- $CC = 0.6$



DEGREE CORRELATION

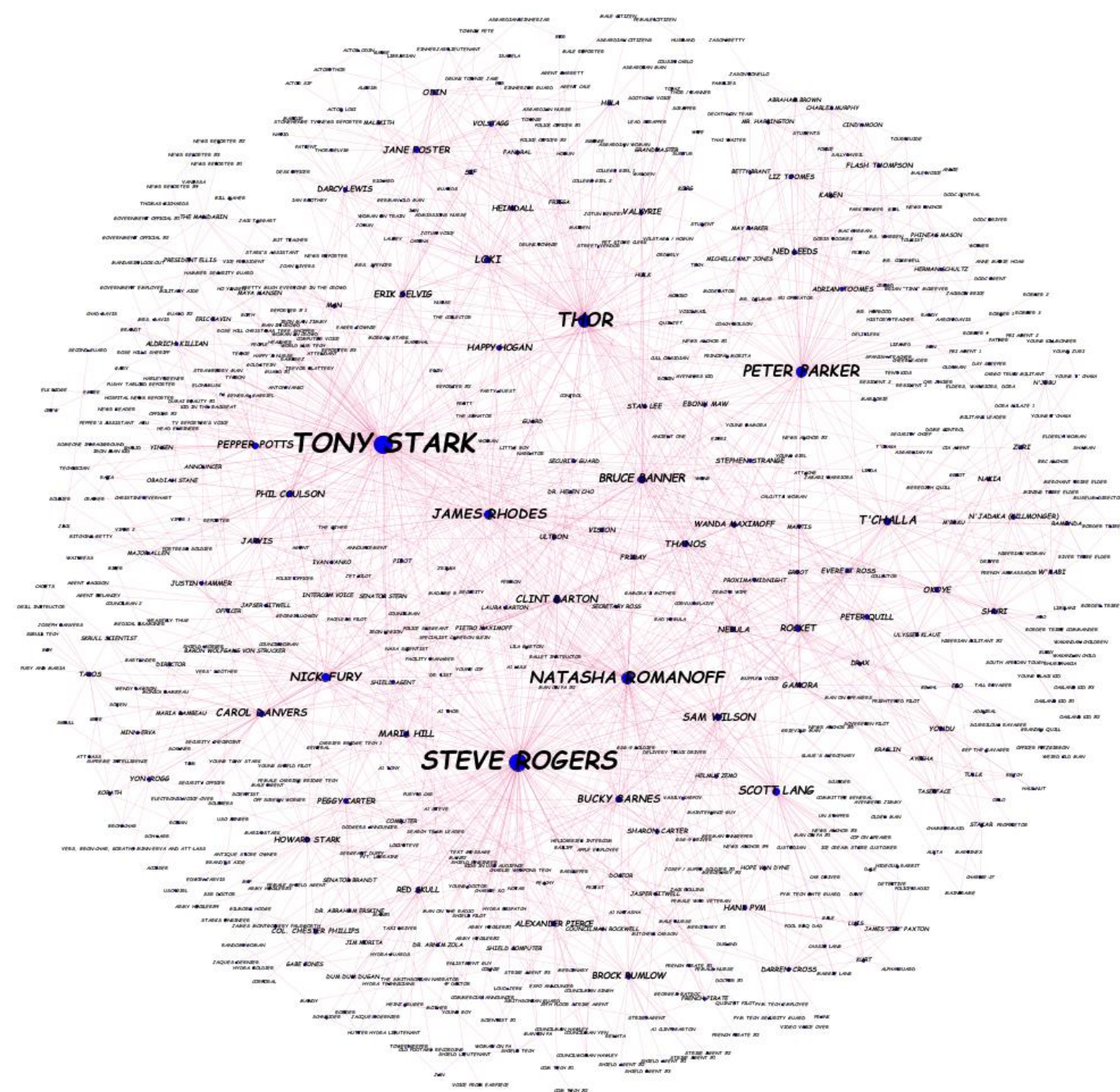
- Capture the relationship between the degrees of nodes that link to each other.
- Assortative network: hubs tend to connect to other hubs.
- Disassortative network: hubs prefer to link to low-degree nodes.
- Annd: the average degree of its neighbor for each node.
- Pearson correlation: annd depends linearly on k with slope r
- If $r > 0$ disassortative, if $r < 0$ assortative, here $r = -0.05$

Average nearest neighbors degree as a function of the degree number



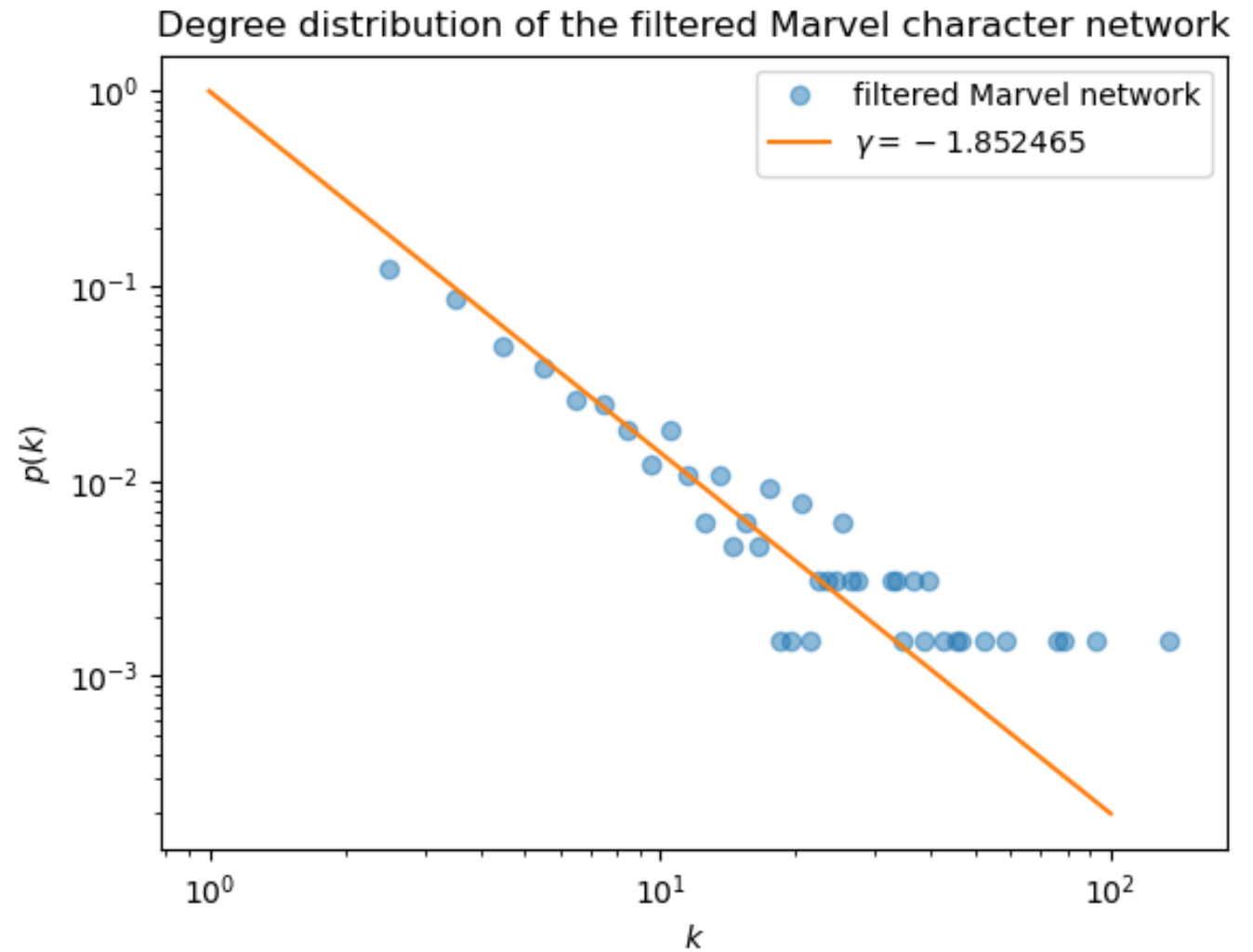
IMPROVING THE NETWORK

- Considering only characters who have a line before or after a certain character, which can be considered a conversation.
- Help us filter out the smaller characters and give us a truer picture of the network.



PROPERTIES

- $\langle k \rangle = 6.22$
- $\langle C \rangle = 0.92$
- $CC = 0.16$
- $r = -0.15$





THE END

Thanks for the attention!